

## University of Groningen

### Misclassification and the use of register-based indicators for depression

Thielen, K.; Nygaard, E.; Andersen, I.; Rugulies, R.; Heinesen, E.; Bech, P.; Diderichsen, F.; Bultmann, U.

*Published in:*  
Acta Psychiatrica Scandinavica

*DOI:*  
[10.1111/j.1600-0447.2008.01282.x](https://doi.org/10.1111/j.1600-0447.2008.01282.x)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2009

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Thielen, K., Nygaard, E., Andersen, I., Rugulies, R., Heinesen, E., Bech, P., Diderichsen, F., & Bultmann, U. (2009). Misclassification and the use of register-based indicators for depression. *Acta Psychiatrica Scandinavica*, 119(4), 312-319. <https://doi.org/10.1111/j.1600-0447.2008.01282.x>

#### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Misclassification and the use of register-based indicators for depression

Thielen K, Nygaard E, Andersen I, Rugulies R, Heinesen E, Bech P, Bültmann U, Diderichsen F. Misclassification and the use of register-based indicators for depression.

**Objective:** To study the degree to which depression indicators based on register data on hospital and antidepressant treatment suffer from differential misclassification with respect to gender, age and social group.

**Method:** Data on 7378 persons were obtained by linking a cross-sectional survey of Danish adults aged 40 and 50 years with population-based registers. Misclassification was analysed by comparing survey data to register data on major depression using the method proposed by Rothman and Greenland.

**Results:** Differential misclassification was found. Adjustment for misclassification reduced women's odds ratios from 2.18 to 1.00 for hospital treatment and from 1.70 to 1.10 for antidepressants. For the lower social group, the corresponding odds ratios increased from 1.18 to 3.52, and from 1.35 to 2.32 respectively, whereas odds ratios with respect to age remained almost unchanged.

**Conclusion:** Differential misclassification should be considered when register-based information about hospital and antidepressant treatment are used as depression indicators.

**K. Thielen<sup>1</sup>, E. Nygaard<sup>1</sup>,  
I. Andersen<sup>1</sup>, R. Rugulies<sup>2</sup>,  
E. Heinesen<sup>3</sup>, P. Bech<sup>4</sup>,  
U. Bültmann<sup>5</sup>, F. Diderichsen<sup>1</sup>**

<sup>1</sup>Section of Social Medicine, Department of Public Health at the University of Copenhagen, Copenhagen, Denmark, <sup>2</sup>National Research Centre for the Working Environment, Copenhagen, Denmark, <sup>3</sup>AKF, Danish Institute of Governmental Research, Copenhagen, Denmark, <sup>4</sup>Psychiatric Research Unit CCMH, Frederiksberg General Hospital, Hillerød, Denmark and <sup>5</sup>Section of Social Medicine, Department of Health Science at the University Medical Center Groningen, the Netherlands

Key words: outcome measurement error; depression; antidepressive agents; population registries

Karsten Thielen, Section of Social Medicine, Department of Public Health at the University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen K, Denmark. E-mail: k.thielen@socmed.ku.dk

Accepted for publication September 19, 2008

## Significant outcomes

- As a result of differential misclassification, gender differences in depression may be substantially overestimated, and social inequality in depression may be substantially underestimated, if hospital treatment and antidepressant treatment are used as indicators for depression.
- Differential misclassification is less important with respect to age, when comparing age groups of 40 and 50 years.

## Limitations

- The Major Depression Inventory, used as 'gold standard', may be misclassified itself compared to clinical assessment. The extent and direction of this misclassification can also be differential with respect to the determinants investigated in this study.
- The results from the Major Depression Inventory are not independent from hospital and antidepressant treatment.
- The results, especially for hospital treatment, are limited due to the small number of cases.

## Introduction

During the last decade, interest in depression as a major public health research issue has constantly grown in the western world (1). Depression rating

scales are the preferred instruments to measure depression in surveys, but their use is limited due to selection bias with respect to non-respondents, lack of statistical power and misclassification. Some recent studies have used population-based register

information, such as hospitalizations with a diagnosis of depression or prescription of antidepressants, as indicators for depression (2–8). These indicators address the first two limitations, but they may introduce misclassification. The misclassification may be differential, as many cases remain untreated and many of the treated cases are selected according to major determinants.

In epidemiological studies on depression, disease status has to be defined and measured for every person. The conventional clinical procedure is to interview a person by a trained person and check for the occurrence and duration of a predefined set of symptoms, which form the basis for the diagnosis according to the 'International Classification of Diseases' (ICD) or 'Diagnostic and Statistic Manual of Mental disorders' (DSM) (9). For epidemiological purposes, this procedure can be used in small-scale studies, but it is often too costly for larger samples. In surveys, one possibility is to ask directly whether a person has had a depression in the past or is currently diagnosed with depression. This information is viewed as an indicator of the life-time prevalence of depression in the survey population. It is more sophisticated to find out whether the individual has a clinical depression at the time of the survey. The key is to ask about a set of depressive symptoms that correspond to the diagnostic schemes and algorithms used in the ICD or DSM. Such rating scales have been used as case-finding tools in clinical trials and epidemiological studies (10, 11). The rising interest in depression research and the mentioned limitations of the rating scales have broadened the view on alternative ways to measure depression as an outcome.

In Denmark, administrative registers are available that cover the whole population and include information about all in- and out-patient hospital contacts as well as all dispensed drug prescriptions. All Danish residents have the right to free medical consultation from a general practitioner (GP), medical consultation from specialists on referral from a GP, free hospital treatment and subsidies for prescribed drugs. Aside from this, residents can directly consult private psychologists or psychiatrists at their own expense. Through data linkage with other registers or survey samples, it is possible to use this register-based information on hospital treatment for depression or on the prescription and dispensation of antidepressants as outcome measures in large-scale longitudinal epidemiological studies, with several measurement points during the follow-up. However, there may be general concerns about the validity of registers (12) as well as special concerns about potential misclassification, which is the focus of this article.

Choosing measurements to define cases and non-cases of depression, one pays particular attention to low sensitivity, because this kind of misclassification means a loss of depression cases. There is further interest in low specificity because this kind of misclassification increases the rate of false positives among those assigned as being depressed. Often, one tends to assume that, in terms of sensitivity and specificity, the degree of misclassification in the outcome is similar or non-differential in all exposure groups, and the resulting odds ratios (ORs) or risk ratios are then biased towards the null. In the case of differential misclassification, however, where the sensitivity or specificity varies across different exposure levels, the results might be seriously biased in either directions (13).

### Aims of the study

The aim of this study is to illustrate the importance of the issue by examining to what degree register data on hospitalization, as a result of depression and antidepressant treatment, are prone to differential misclassification with respect to gender, age and social group. Comparing the occurrence of depression measured by a rating scale in a survey with the register data on depression treatment, the sensitivity and specificity will be calculated for gender, age and social groups. These results will be used to adjust the crude associations between determinants and hospital and antidepressant treatment for differential misclassification.

### Material and methods

#### Study population and variables

Data were obtained by linking a cross-sectional survey of Danish adults aged 40 and 50 years, carried out by March 2000, with three population-based administrative registers. The linkage procedure makes use of the fact that each resident of Denmark is assigned a unique central person identification number. For research purposes, Statistics Denmark is able to combine survey samples with register information at the individual level through an anonymous process, as done for this study.

The following data sources have been used:

- i) The survey sample is part of the 'Danish Longitudinal Study on Work, Unemployment and Health', which contains two representative samples of 40- and 50-year-old persons and a supplemental sample of marginalized persons, which was not included in this study. Response rates were 68% and 69%

respectively for the two-age samples (14). Data on non-participants showed a higher proportion of men, immigrants, and semiskilled or unskilled workers as well as a higher proportion of hospital treatment for depression and antidepressant treatment. Depression was measured with the 'Major Depression Inventory' (MDI). The inventory consists of 10 items that cover the DSM-IV diagnosis of major depression. It assesses information on depressive symptoms with a continuous duration of at least 2 weeks, which corresponds to DSM-IV and ICD-10. The MDI has been validated at clinical and population levels (10, 15, 16) with respect to the DSM-IV diagnosis. Moreover, information on gender and age was taken from the survey material.

- ii) The Danish Psychiatric Central Register (17) includes all persons in Denmark who have been treated by the psychiatric hospital system since 1970. From the register, we derived the cases with the main diagnosis of depression at admission, including bipolar affective disorder (ICD-10: F31, F32, F33 and F34), during the time period between 1998 and 2002. Data on all patients who had been treated as in- and out-patients, but not single contacts with the hospital ambulatory or emergency department, were used. The admission prevalences in 2000 and for the period from 1998 to 2002 were chosen as prevalence outcome measures. Hospital admissions during the year 2000, without registrations in the preceding 4 years, were defined as incident cases. For power reasons, prevalence of depression from 1998–2002 was used for the stratified analysis of sensitivity and specificity.
- iii) The Medicinal Product Statistics (18) contain data on all prescribed medication, which has been dispensed at pharmacies in Denmark since 1995. There is no other legal way in Denmark to obtain prescribed medication (except from hospitals in connection with admission). The registrations are coded according to the Anatomical Therapeutic Chemical (ATC) classification system (19). Registrations for all antidepressants (ATC: N06A), were used. In addition to the ATC codes, information about the date of prescription and the defined daily dose (DDD) was taken. Prevalence variables as indicators for antidepressant treatment were defined for the year 2000 and for the period from 1998 to 2002. In the same way as for the hospital admissions, a measure for the incidence in 2000 was defined. Finally, in order to gain

enough power for the analysis of sensitivity and specificity, we used a measure for the 3-year cumulative incidence (from 2000 to 2002). With respect to the cumulative amount of DDD, we used two measures: a) the dispensing of at least one DDD of any antidepressant, and b) more than 179 DDD of any combination of different antidepressants, which indicates a continuous treatment, for at least 6 months.

- iv) Because of the high non-response rate with respect to questions about social position in the survey material, administrative register information from Statistics Denmark at the 'level of education' obtained before the year 2000 as the indicator for social position was taken. To increase the power of the analysis, we dichotomized it into two social groups: 'low' and 'middle/high' where low means no education beyond lower secondary school.

As a result of the data-linkage procedure, we were able to analyse a sample of 7378 persons who participated in the survey and had information on all variables used in the analysis.

#### Statistical analysis

In the initial analysis, the prevalence and incidence of depression, hospital treatment for depression and antidepressant treatment for the sample of 40- and 50-year-old persons was calculated. The crude ORs for prevalent hospital treatment and prevalent and incident antidepressant treatment with respect to gender, age and social group were also calculated.

It is a central assumption in this analysis that the MDI measure of depression provides the most reliable information about a person's state of depression. We used the algorithm for 'Major Depression' only, which follows the DSM-IV rules to classify persons as depression cases, because there were not enough cases to work with different levels of depressiveness from the ICD-10. As measures of the degree of accordance between MDI and hospital treatment, and MDI and prescription and delivery of antidepressants, the sensitivity (percentage of persons indexed as major depression cases who had hospital or antidepressant treatment respectively) and the specificity (percentage of persons indexed as non-cases of major depression who have no hospital or antidepressant treatment respectively) were calculated for every stratum of gender, age and social group. Sensitivity and specificity can be similar or different across the strata, which was assumed to be a sign

of non-differential or differential misclassification respectively. In the final analysis, the former calculated crude ORs for misclassification was adjusted with respect to the results for sensitivity and specificity in the different strata. In this exemplary analysis, the outcome measure was adjusted using a method based on sensitivity (Se), specificity (Sp) and false positive (Fp), which has been described by Rothman and Greenland (20).

The following formulae were used:

$$A = \frac{(A^* - Fpn)}{(Se + Sp - 1)} \quad \text{and} \quad B = n - A,$$

where  $A$  and  $B$  represent the estimated 'true' numbers of diseased and non-diseased subjects respectively and  $A^*$  the number of subjects classified as diseased by the measure under study. These equations are used for every stratum respectively. New adjusted estimates can then be calculated from these 'corrected' counts, which have been done for the ORs in this study.

## Results

Table 1 shows the descriptive results from the study. For the year 2000, the estimated point prevalence for 'Major Depression' in the age groups of 40 and 50 years was about 2.7%. During the year 2000, about 5% of the population received some kind of antidepressant treatment at least one time. More than every second person of this former group and 2.8% of the population sample received at least 180 DDD, which is equivalent to 6 months medical treatment. About 1.2% of the population started antidepressant treatment during the year 2000, after at least

Table 1. Indicators of depression and their prevalence and incidence in a sample of the Danish population, percentages and absolute numbers for 2000, 1998–2002 and the incidence for 2000 ( $n = 7378$ )

Indicators for depression	Prevalence 2000	Prevalence 1998–2002	Incidence 2000*
Major Depression Inventory, DSM-IV algorithm	2.7 (196)		
In- and out-patients†	0.18 (13)	0.83 (63)	0.11 (8)
In-patients†	0.09 (7)	0.43 (32)	0.04 (3)
Antidepressants, $\sum DDD > 0\ddagger$	5.04 (372)	10.00 (738)	1.17 (86)
Antidepressants, $\sum DDD > 6$ months‡	2.81 (207)	6.09 (449)	0.33 (24)

DDD, defined daily dose.

\*No registrations in the preceding 4 years.

†Hospital register, diagnose F31–F34 (ICD-10).

‡Medicinal Product Statistics.

4 years without treatment. Among those, 28% received at least 180 DDD.

The point prevalence for depression according to hospital treatment was 0.2% during the year 2000 and 0.8% during the 5-year period between 1998 and 2002.

As shown in Table 2, compared to men, women had up to twice as high OR for hospital and antidepressant treatment, while the OR with regard to the MDI was similar in men and women. Due to a lack of power, we calculated the OR for the 3-year cumulative incidence (2000–2002) only for prescribed antidepressants. The OR was again higher for women than for men, but less pronounced than ORs for the prevalence, indicating a longer duration of antidepressant treatment for women.

The older age group had higher odds for hospital treatment (not significant) and antidepressant treatment, independent of the time frame. As was seen for gender, the OR for the cumulative incidence of antidepressant use was at a lower level than for prevalence, but not significant.

Table 2. Odds ratios and 95% confidence intervals for the indicators of prevalent depression in 2000 and 1998–2002, and the cumulative incidence between 2000 and 2002 by gender, age and social group

	Prevalence 2000 OR (95% CI)	Prevalence 1998–2002 OR (95% CI)	Cumulative incidence 2000–02* OR (95% CI)
Gender (female)			
Major Depression Inventory, DSM-IV algorithm	1.01 (0.76–1.35)	–	–
In- and out-patients†	2.04 (0.63–6.64)	2.18 (1.25–3.78)	–
Antidepressants, $\sum DDD > 0\ddagger$	1.78 (1.43–2.21)	1.70 (1.46–2.00)	1.45 (1.14–1.84)
Antidepressants, $\sum DDD > 6$ months‡	1.76 (1.32–2.36)	1.71 (1.40–2.08)	1.32 (0.83–2.09)
Age (50 years)			
Major Depression Inventory, DSM-IV algorithm	1.12 (0.84–1.48)	–	–
In- and out-patients†	2.25 (0.69–7.30)	1.26 (0.76–2.09)	–
Antidepressants, $\sum DDD > 0\ddagger$	1.47 (1.19–1.81)	1.35 (1.16–1.57)	1.13 (0.90–1.43)
Antidepressants, $\sum DDD > 6$ months‡	1.60 (1.21–2.13)	1.54 (1.27–1.88)	1.30 (0.83–2.05)
Social group (low)			
Major Depression Inventory, DSM-IV algorithm	2.34 (1.75–3.12)	–	–
In- and out-patients†	0.56 (0.12–2.51)	1.18 (0.68–2.08)	–
Antidepressants, $\sum DDD > 0\ddagger$	1.40 (1.11–1.77)	1.35 (1.14–1.60)	1.10 (0.84–1.43)
Antidepressants, $\sum DDD > 6$ months‡	1.35 (1.00–1.83)	1.30 (1.05–1.60)	1.02 (0.61–1.72)

OR, Odds ratio; 95% CI, 95% confidence interval; DDD, defined daily dose.

\*No registrations in the preceding 4 years.

†Hospital register, diagnose F31–F34 (ICD-10).

‡Medicinal Product Statistics.



Table 3. Comparison of cases of 'major depression' measured with the MDI in 2000 and depression indicators from the registers

Gender	Se	Sp	Se	Sp
	Male		Female	
In- and out-patients, 1998–2002*	4.3 (1.2–10.7)	99.6 (99.3–99.8)	11.5 (6.1–19.3)	99.2 (98.3–99.4)
Antidepressants, 2000†	30.1 (21.0–40.5)	97.1 (96.5–97.6)	33.7 (24.7–43.6)	94.5 (93.7–95.2)
Antidepressants, 1998–2002‡	44.1 (33.8–54.8)	93.4 (92.6–94.3)	54.8 (44.7–64.6)	89.0 (87.9–89.9)
Antidepressants, Inc., 2000–2002‡	10.8 (5.3–18.9)	96.9 (96.3–97.5)	16.4 (9.8–24.9)	95.6 (94.9–96.3)
Age	40 years		50 years	
In- and out-patients, 1998–2002*	9.7 (4.5–17.6)	99.5 (99.2–99.7)	6.7 (2.8–13.4)	99.3 (98.9–99.5)
Antidepressants, 2000†	33.3 (23.9–43.9)	96.7 (96.0–97.2)	30.8 (22.1–40.6)	94.7 (94.0–95.4)
Antidepressants, 1998–2002‡	50.5 (40.0–61.1)	92.4 (91.5–93.3)	49.0 (39.1–59.0)	89.8 (88.7–90.7)
Antidepressants, Inc., 2000–2002‡	11.8 (6.1–20.2)	96.4 (95.8–97.0)	15.4 (9.1–23.8)	96.1 (95.4–96.7)
Social group	Low		Middle/high	
In- and out-patients, 1998–2002*	9.5 (4.2–17.9)	99.5 (99.0–99.8)	7.1 (3.1–13.5)	99.3 (99.1–99.5)
Antidepressants, 2000†	36.9 (26.6–48.1)	95.1 (94.0–96.1)	28.3 (20.2–37.6)	95.9 (95.3–96.4)
Antidepressants, 1998–2002‡	46.4 (35.5–57.7)	89.5 (88.0–90.9)	52.2 (42.6–61.7)	91.6 (90.8–92.3)
Antidepressants, Inc., 2000–2002‡	9.5 (4.2–17.9)	96.0 (94.9–96.8)	16.8 (10.4–25.0)	96.4 (95.8–96.8)

Se, sensitivity; Sp, specificity; DDD, defined daily dose.

\*Hospital register, diagnoses F31–F34 (ICD-10).

†Medicinal Product Statistics, antidepressants  $\sum$  DDD > 0.

‡Medicinal Product Statistics, antidepressants  $\sum$  DDD > 0, cumulative incidence, no registrations in the preceding 4 years.

For the point and period prevalence, we found that the lower social group has a 30–40% higher odds for antidepressant treatment; for the incidence that no significant difference was found between the social groups. The OR for MDI as the outcome was 2.34 in the lower social group, which was more than twice as high odds as the middle/high social group.

Table 3 shows the sensitivity and specificity for the register data on hospital and antidepressant treatment by gender, age and social group. Hospital treatment as an indicator showed a sensitivity below 10%, which means that 90% of the persons categorized as depressive by the MDI in 2000 did not have any hospital treatment between 1998 and 2002. On the other hand, hospital treatment showed the highest specificity for all groups, around 99.5%. Thus, only about 0.5% among those not treated by the psychiatric hospital system between 1998 and 2002 has been identified by the MDI as a depression case.

When comparing men and women, we found large, but non-significant, differences in sensitivity for hospital and antidepressant treatment in 2000, during 1998–2002 and for the 3-year cumulative incidence. The specificity for antidepressant treatment in 2000 and during 1998–2002, however, was significantly different, which indicates differential misclassification with respect to gender. The differences between the two-age groups for the sensitivity were prominent, but non-significant for the prevalence of hospital treatment and the incidence of antidepressant treatment. For the prevalence of antidepressant treatment, we found

significantly higher specificity for the younger age group indicating differential misclassification with respect to age.

For the social groups, there were marked, but non-significant, differences in sensitivity for all measures, but only small and insignificant differences in specificity between the indicators of depression. From this, we cannot exclude differential misclassification with respect to social group.

Finally, we adjusted the ORs for hospital and antidepressant treatment by using the calculated estimates for sensitivity and specificity (Table 4). While the crude prevalence odds for women were

Table 4. Odds ratio for the indicators of depression in 2000, 1998–2002 and the cumulative incidence between 2000 and 2002 by gender, age and social group, adjusted for differential misclassification in the outcome

	Prevalence 2000		Prevalence 1998–2002		Cumulative incidence 2000–2002	
	Crude OR	Adjusted OR	Crude OR	Adjusted OR	Crude OR	Adjusted OR
Gender (female)						
In- and out-patients*	–	–	2.18	1.00	–	–
Antidepressants†	1.78	1.03	1.70	1.10	1.45	1.37
Age (50 years)						
In- and out-patients*	–	–	1.26	1.50	–	–
Antidepressants†	1.47	1.01	1.35	1.20	1.13	1.14
Social group (low)						
In- and out-patients*	–	–	1.18	3.52	–	–
Antidepressants†	1.40	2.28	1.35	2.32	1.10	2.27

OR, Odds ratio; DDD, defined daily dose.

\*Hospital register, diagnoses F31–F34 (ICD-10).

†Medicinal Product Statistics, antidepressants  $\sum$  DDD > 0.

up to twice as large as the odds for men, the adjusted ORs came close to 1. On the contrary, the gender OR for the incidence of prescription and delivery of antidepressant treatment remained almost unchanged. The age difference, with higher odds for the older age group for antidepressant treatment, disappeared for the point prevalence in 2000, but not for the period prevalence from 1998–2002. The age OR for incidence of antidepressant prescription remained almost unchanged, but the OR for prevalence of hospital treatment increased. The largest changes were observed for the social groups where all ORs increased substantially after adjustment.

## Discussion

The study demonstrated a considerable degree of differential misclassification when we used the prevalence of hospital treatment or the prevalence and incidence of antidepressant treatment as outcome indicators for depression compared with the major depression rating scale. Adjustment of the crude effect estimates in the analyses showed that one would strongly overestimate the gender effect on depression (positive bias in the odds for females compared to males) and underestimate the effect of social group on depression (negative bias in the OR of low social group compared to the middle/high group), when the prevalence of hospital or antidepressant treatment were used as indicators. The use of incident antidepressant treatment as an indicator gave a fairly unbiased OR estimate for gender. The effect of social group, however, is negatively biased. A consistent pattern in the results of the adjustment, which would favour one of the tested indicators in general, was not found.

The higher risk of antidepressant treatment for the lower social group confirms another Danish study (21), but differs from a Finish study, which finds antidepressant treatment less common among low educational groups (3). Differential and non-differential misclassification of the measure of outcome, but also of the exposure, is a well-studied source of bias (22–24). Non-differential misclassification in epidemiological studies can be dealt with because we know it changes the effect estimate towards null, which results in an underestimation of the true association. Several correction methods are available (25, 26). Differential misclassification, however, gives rise to more concern, because bias may go in either direction, as illustrated in the present study. This should be accounted for when register information is used, even when data quality is high, as in the Danish case with a

population coverage of about 100% and good diagnostic validity for the hospital-treated cases (27). Another important issue is that epidemiological studies have to deal with misclassification not only for the outcome, but also for the exposure as well as for all possible confounders. When several misclassification problems have to be dealt with at the same time, there is a need for more advanced correction methods (25, 26). These potential additional sources of bias might affect the risk estimate substantially in both directions.

The low sensitivity for hospital treatment, also within a 5-year time period, reflects that only a very small part of the persons who participated in the survey and classified as depressed by the MDI received treatment in the hospital system. This problem is even amplified by selection bias. In general as in this population study, more severe cases do not participate in surveys, and at the same time are more likely to be found in hospital registers. Because non-response rates are different for gender, age and social groups, this could also differentially affect the rates and ratios, but because we expect the MDI to measure with similar precision among non-participants, it would not affect the adjustment.

Healthcare utilization bias is also relevant with regard to antidepressant treatment; not everybody with a depression is getting antidepressants, because there are other treatment options, patients do not want them or there is no contact with the healthcare system. Again this can affect the exposures differently. For example, may higher social groups have more frequently access to and use psychotherapy compared to lower social groups, which would result in differential misclassification. Also, there are other medical indications besides depression for which antidepressants are a treatment choice. Anxiety, pain and sleep disturbances are the most frequent (28, 29). This might increase the number of false-positive cases in the analysis, which in turn affects the specificity. Gender, age or social group differences in prevalence of those indications for antidepressant treatment might again result in differential misclassification. Moreover, missing true cases and false-positive cases may result from measurement error with respect to the MDI. A population-based validation study (10) finds misclassification of the MDI measure, expressed in sensitivity (67%) and specificity (81%); unfortunately, we do not know the corresponding values for this study sample. If we rely on these validation results, we could hardly use the major depression prevalence measured by MDI in absolute terms. As we use only

relative measures and an adjustment method that is independent of the prevalence, however, we think it is acceptable for this exemplifying analysis. On the other hand, we cannot rule out differential misclassification with respect to gender, age or social group for the measurement of the MDI, which is a potential bias to the study results.

Major Depression Inventory defines persons as cases of major depression if they have been depressed according to the DSM-IV algorithm for at least 2 weeks preceding the survey. This point prevalence can, by definition, only capture a certain amount of all persons who had a major depression at least once in 2000 or between 1998 and 2002. Repeated measurement of the MDI would provide a better estimate of the 1-year prevalence for major depression, as it improves the chances of finding cases of depression at other times of the year and which accordingly would be indexed in this analysis as false positives. Assuming that all exposure groups had similar degrees of misclassification at other measurement points during the year, we do not expect this to cause bias in the adjustment. More measurement points, however, would improve precision. Generally, this point prevalence measurement issue results in a more conservative estimate of the specificity.

Yet another important issue relates to the fact that we use treatment measures for depression, which have a potential effect on the state of depression and on the MDI score respectively. If the treatment takes place around the time of the survey, one would expect successfully treated persons not to be indexed as major depression cases. This again would result in a lower specificity. If treatment response on MDI is differential in relation to gender, age and social group, this would affect the results, but we cannot account for this in this study. Finally, one has to keep in mind that gender, age and social position are often included as covariates in epidemiological studies. Differential misclassification may also in these situations affect the results when depression indicators are the outcome.

### Acknowledgements

The study was supported by grants of the Danish Work Environment Fund (grant numbers: 24-2005-09 and 2-2006-04), and of the Ministry of Health and Prevention, Public Health Fund (grant number: 2005-14033-8). The funding agencies were not involved in any stage of the study, including data collection and analysis, manuscript writing and decision to submit the manuscript for publication. All researchers are independent from the funding agencies.

### Declaration of interest

None.

### References

1. PRINCE M, PATEL V, SAXENA S et al. No health without mental health. *Lancet* 2007;**8**:370.
2. DALTON SO, MELLEMKJAER L, OLSEN JH, MORTENSEN PB, JOHANSEN C. Depression and cancer risk: a register-based study of patients hospitalized with affective disorders, Denmark, 1969–1993. *Am J Epidemiol* 2002;**155**:1088–1095.
3. KIVIMAKI M, GUNNELL D, LAWLOR DA et al. Social inequalities in antidepressant treatment and mortality: a longitudinal register study. *Psychol Med* 2007;**37**:373–382.
4. LAWLOR DA, HART CL, HOLE DJ, GUNNELL D, DAVEY SG. Body mass index in middle life and future risk of hospital admission for psychoses or depression: findings from the Renfrew/Paisley study. *Psychol Med* 2007;**37**:1151–1161.
5. VIRTANEN M, HONKONEN T, KIVIMAKI M et al. Work stress, mental health and antidepressant medication findings from the Health 2000 Study. *J Affect Disord* 2007;**98**:189–197.
6. VIRTANEN M, KIVIMAKI M, FERRIE JE et al. Temporary employment and antidepressant medication: a register linkage study. *J Psychiatr Res* 2008;**42**:221–229.
7. WIECLAW J, AGERBO E, MORTENSEN PB, BONDE JP. Occupational risk of affective and stress-related disorders in the Danish workforce. *Scand J Work Environ Health* 2005;**31**:343–351.
8. WIECLAW J, AGERBO E, MORTENSEN PB, BURR H, TUCHSEN F, BONDE JP. Work related violence and threats and the risk of depression and stress disorders. *J Epidemiol Community Health* 2006;**60**:771–775.
9. JABLENSKY A. The nature of psychiatric classification: issues beyond ICD-10 and DSM-IV. *Aust NZ J Psychiatry* 1999;**33**:137–144.
10. FORSELL Y. The Major Depression Inventory versus Schedules for Clinical Assessment in Neuropsychiatry in a population sample. *Soc Psychiatry Psychiatr Epidemiol* 2005;**40**:209–213.
11. OLSEN LR, MORTENSEN EL, BECH P. Prevalence of major depression and stress indicators in the Danish general population. *Acta Psychiatr Scand* 2004;**109**:96–103.
12. BYRNE N, REGAN C, HOWARD L. Administrative registers in psychiatric research: a systematic review of validity studies. *Acta Psychiatr Scand* 2005;**112**:409–414.
13. GREENLAND S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol* 1996;**25**:1107–1116.
14. CHRISTENSEN U, LUND R, DAMSGAARD MT et al. Cynical hostility, socioeconomic position, health behaviors, and symptom load: a cross-sectional analysis in a Danish population-based study. *Psychosom Med* 2004;**66**:572–577.
15. BECH P, RASMUSSEN NA, OLSEN LR, NOERHOLM V, ABILDGAARD W. The sensitivity and specificity of the Major Depression Inventory, using the Present State Examination as the index of diagnostic validity. *J Affect Disord* 2001;**66**:159–164.
16. OLSEN LR, JENSEN DV, NOERHOLM V, MARTINY K, BECH P. The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. *Psychol Med* 2003;**33**:351–356.
17. MUNK-JORGENSEN P, KASTRUP M, MORTENSEN PB. The Danish psychiatric register as a tool in epidemiology. *Acta Psychiatr Scand Suppl* 1993;**370**:27–32.



18. DANISH MEDICINES AGENCY (LÆGEMIDDELSTYRELSEN). Danish products statistics. Available at: <http://www.laegemiddelstyrelsen.dk> (cited 10 January 2008).
19. WHO COLLABORATING CENTRE FOR DRUG STATISTICS METHODOLOGY. Guidelines for ATC classification and DDD assignment. Oslo: WHO Collaborating Centre for Drug Statistics, 2004.
20. ROTHMANN K, GREENLAND S. Modern epidemiology, 3rd edn. Philadelphia, PA: Lippincott-Raven, 2006.
21. HANSEN DG, SONDERGAARD J, VACH W et al. Socio-economic inequalities in first-time use of antidepressants: a population-based study. *Eur J Clin Pharmacol* 2004;**60**:51–55.
22. COPELAND KT, CHECKOWAY H, McMICHAEL AJ, HOLBROOK RH. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977;**105**:488–495.
23. HOFER M. The effect of misclassification on the estimation of association: a review. *Int J Methods Psychiatr Res* 2005;**14**:92–101.
24. WALTER SD, IRWIG LM. Estimation of test error rates, disease prevalence and relative risk from misclassified data: a review. *J Clin Epidemiol* 1988;**41**:923–937.
25. BASHIR SA, DUFFY SW. The correction of risk estimates for measurement error. *Ann Epidemiol* 1997;**7**:154–164.
26. GUSTAFSON P. Measurement error and misclassification in statistics and epidemiology, 1st edn. Boca Raton: Chapman & Hall/CRC, 2004.
27. KESSING LV. Validity of diagnoses and other clinical register data in patients with affective disorder. *Eur Psychiatry* 1998;**13**:392–398.
28. ANGST J. Psychiatric diagnoses: the weak component of modern research. *World Psychiatry* 2007;**6**:30–31.
29. TRIFIRO G, BARBUI C, SPINA E et al. Antidepressant drugs: prevalence, incidence and indication of use in general practice of Southern Italy during the years 2003–2004. *Pharmacoepidemiol Drug Saf* 2007;**16**:552–559.